数据期刊的政策调研与分析
——以 *Scientific Data* 为例

■刘晶晶[1,2)]　顾立平[1)]

收稿日期:2015-01-21
修回日期:2015-03-04

1）中国科学院文献情报中心,北京市中关村北四环西路 33 号　100190
2）中国科学院大学,北京市石景山区玉泉路 19 号(甲)　100049

摘　要　【目的】探索数据期刊良好实践的实施细节,对其政策要素框架进行分析,以期为我国数据期刊的发展提供参考。【方法】根据实用主义方法论,以文献综述的方式进行分析,以合理竞替的方法筛选和提炼政策要素。【结果】数据期刊应该着重数据期刊提交规范(提交对象、内容格式要求、确保数据真实可查证)、质量管理流程(数据描述符内容信息、数据描述符产生流程、数据知识库的审核)以及保障各方权益的声明(读者的权益、作者的权益、科研机构与资助机构等的权益)。【结论】在权威期刊的示范效应下,制定符合我国的数据期刊发展政策,可推动科研数据开放共享。

关键词　科研数据;数据共享;数据期刊;*Scientific Data*
DOI:10. 11946/cjstp. 201501210077

1　科学数据共享的大趋势

科学研究的数字化、网络化和开放化特点,促进了全球科研社群对于海量数据共享与利用的需求。为此,科学家们在 2003 年的《关于自然科学与人文科学资源的开放获取的柏林宣言》(Berlin Declaration on Open Access to Knowledge in the Science and Humanities)中,明确将科研数据作为学术知识的一部分[1]。联合国教科文组织(United Nations Educational, Scientific and Cultural Organization,简称 UNESCO)也将科学数据界定为三种开放获取信息资源(科学出版物、教育资源、科学数据)之一[2]。在 2007 年经济合作发展组织(Organization for Economic Co-Operation and Development,简称 OECD)发布的《公共资金科研数据获取的原则与指南》中,提出“促进科研人员之间的数据访问和共享”作为宗旨[3]。2012 年英国皇家学会发布报告《科学是开放事业》[4]中也认为,未来可以作为真正能够可利用的科学数据是以智能化开放数据(intelligently open data)为表现形式,同时“敦促科研人员的科研数据可获取和可利用,科学家们应该在合适的数据知识库里存储数据,让人使用和验证数据”[5]。由此可见,科研数据的开放共享已经成为一种新兴的学术要求和学术态势。

2　数据期刊的诞生及其发展

为了确保科研数据和相关信息可获取、可理解和可使用,科研教育机构、资助机构、科研人员、图书馆员、数据中心与数据知识库管理员、学术委员会等都发挥了重要的贡献[6]。科学数据开放共享的实践大致分为三种类型,归纳如图 1 所示:

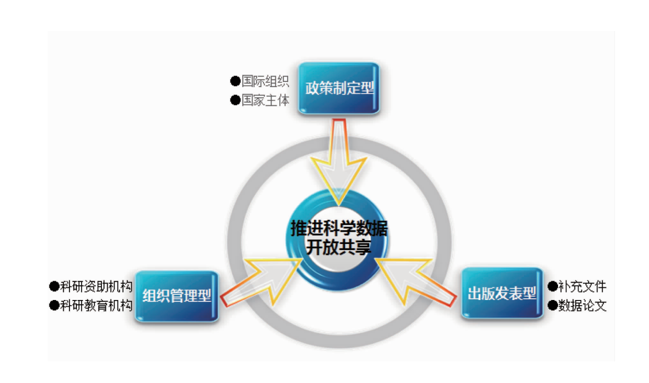


图 1　推进科学数据开放共享的三种实践类型

基金项目:《中国科学院文献出版领域择优支持人才》项目(项目编号:院 1434)
第一作者简介:刘晶晶(ORCID:0000 -0001 -9065 -7345),硕士研究生,E-mail:liujingjing@ mail. las. ac. cn
通讯作者:顾立平(ORCID:0000 -0002 -2284 -3856)博士,副研究员,E-mail:gulp@ mail. las. ac. cn

就出版发表型而言,主要是指科技期刊在数据共享和数据出版方面的探索。一直以来,出版界都将科学成果的再现作为一个高质量学术刊物对于科学界所应担负的责任。过去十年间,STM 出版社对于科学论文和科学数据的引用以及科研人员在这方面的贡献有了长足的进步[7]。国际上有很多期刊都制定了“科学数据共享政策”。这些期刊要求作者在投稿时必须向期刊编辑和同行评审专家提供相关的科学数据或者提供可以获得该研究所涉及科学数据的第三方数据存储库的存取号,如果不能提供,则必须进行说明[8],比如:*Biodiversity Data Journal*[9]、*Ecology*[10]、*Earth System Science Data*[11]等。

随着科技期刊数据共享政策的推进,继而出现了数据期刊和数据论文。以自然出版集团(NPG)为例,在期刊出版政策中制定了严格促进科学数据共享的政策,明确指出“作者必须不设任何限制地提供材料、数据和有关协议给其他人”[12]。在此基础上,2014 年 5 月 NPG 又参考科研资助机构的数据要求,科研人员、图书馆员、数据知识库管理员、数据标准倡议者对于科研数据管理方式的调查结果、相关权益和意见[13],推出了在线出版的开放获取期刊 *Scientific Data*[14](以下简称 *SD*),以文章的形式发布有科学价值的数据描述。

就国内而言,我国也已经建设了一批数据共享中心,如国家农业科学数据共享中心[15]、国家人口与健康科学数据共享平台[16]以及地球系统科学数据共享平台[17]。其中地球系统科学数据共享平台的建设与孙九林院士等科学家持续不断的努力密切相关,有大量的文献以此为案例进行了探讨[18-19]。另外,中国科学院计算机网络信息中心黎建辉教授在 2014 年开放获取周上做报告,从构建科学数据标准规范、流程与出版平台,促进数据开放共享、引用、评价和探索行之有效的可持续数据出版机制三个方面进行了详细的阐释[20],并积极展开相关的实践探索。

数据期刊是作为承认科研人员在提供科研数据上对科学进步产生贡献的一种记录方式。由于它作为“数据期刊”有不同于“学术期刊”的地方,因此,如何界定、开展和经营数据期刊,有赖对它的期刊政策进行分析。现有的主要数据期刊如表 1 所示:

表 1 现有的主要数据期刊

序号	期刊名	出版社	出版社类型	简介
1	*Ecological Archives-Data Papers*[21]	ESA	学会出版	该期刊主要是生态学领域的论文,提供数据论文的长期访问和维护,旨在促进数据的二次使用和分析,并鼓励定期更新数据集。数据论文接受完整的同行评审,但同时也接受进一步技术审查
2	*The Journal of Chemical & Enigineering Data*[22]	ACS	学会出版	致力于出版实验和计算数据,是美国化学学会唯一一本关注于包含数据论文的期刊,关注于物理学、热力学和传播属性定义良好的材料,包括已知成分的复杂混合物
3	*CODATA's Data Science Journal*[23]	CODATA	学会出版	一本同行评议的开放获取网络期刊,发表有关科学数据和数据集管理的文章,覆盖的科学包括:物理科学、工程、地球科学和生物科学以及农业和医学科学
4	*Dataset Papers in Science*[24]	Hindawi publishing	新兴 OA 出版	一本同行评议的开放获取期刊,出版科学和医学方面的论文数据集。该刊是由相关领域的国际编委会专家运营的
5	*Earth System Science Data*[25]	Copernicus	新兴 OA 出版	一本发表文章原始研究数据的国际期刊,旨在发表高质量的地球系统科学原始数据
6	*Ubiquity Press metajournals*[26]	Ubiquity Press	新兴 OA 出版	该期刊提供一种完全开放的方式来发现通常很难找到的分布在多个位置的研究资源,鼓励作者公开存档其研究数据,通过引用跟踪相关报告进而展开新的合作
7	*F1000 Research*[27]	F1000 Research	新兴 OA 出版	一种新的基于开放透明的、发表后同行评审的开放获取及发表模式,主要针对生物学和医学领域

续表 1

序号	期刊名	出版社	出版社类型	简介
8	*GigaScience*[28]	BMC	新兴 OA 出版	采用标准全文文献、数据库信息以及信息分析工具相结合的崭新模式，来发表大规模的生物学研究成果，使读者可以直接通过文献所提供的数据和分析工具对结果进行测试和验证
9	*Geoscience Data Journal*[29]	Wiley	商业出版	为科学数据的正式出版提供了一个开放的平台。发表的一篇论文描述了一个数据集，给出详细的收集、处理、软件、文件格式等，没有文本分析或者创造性结论的要求
10	*Scientific Data*[14]	Nature	商业出版	以文章的类型发布有科学价值的数据描述
11	*Data in Brief*[30]	Elsevier	商业出版	通过出版数据论文，方便科研人员之间共享和重用彼此的数据集，从而有利于科研成果的检验和再现
12	*International Journal of Robotics Research DataPapers*[31]	Sage	商业出版	第一个有关机器人研究的学术出版物，关注人工智能、应用数学、计算机科学、机电工程等领域，同时还出版同行评审的数据论文以及与文章相关联的多媒体延伸
13	*Biodoversity Data Journal*[32]	Pensoft	商业出版	一个开放获取的在线平台，旨在加快出版、传播和共享与生物多样性有关的数据

本文以 *SD* 为例，具体分析数据期刊在提交规范、质量规范和其他相关权益规范方面的政策要素，以期能够在权威期刊的示范效应下，制定符合我国的数据期刊发展政策，推动科研数据开放共享。

3 研究方法与研究框架

本文研究设计的方法参考《前瞻导论》[33]所提到“政策内容采集”到“政策建议提出”的过程，采用政策观测的方式设置政策观测重点，据此设计研究框架，如表 2 所示。在研究过程中参考《型人导论》[34]第二章，以实用主义方法论的四大原则判断具体内容选取与否，并且在不同政策内容中，以合理竞替的方法筛选能够解释现况的最佳选择方案，据此提炼政策要素。形成四、五、六节的论证内容，并在第七节形成本文经过分析后的建议方案。

数据期刊是应对科研数据共享的大趋势之下发展起来的新型学术期刊，是一种对数据的使用（对数据的描述）、引用（对数据的引用）和重用（对数据内容来源的准确识别和溯源）都非常重要的工具。同学术期刊相比，其对象、受众、特点和实质都有明显的差别，整理如表 3 所示。

承上，一般意义上的学术期刊论文是对背景、目的、意义、方法、结论等的验证，而数据期刊的论文则涉及到数据描述，两者在提交内容、出版对象等方面有显著的不同，因而涉及到的规范要求就会有所差异。

表 2　数据期刊政策研究内容

研究问题	政策观测重点（解答研究问题的分析重点）
作者如何提交数据论文？	1. 提交怎样的数据？例如是数据生命周期中的所有数据还是科研共享中的数据。 2. 所提交的数据论文应该怎样结构化和规范化描述数据？例如除了数据的元数据内容，还需要哪些描述数据的元素。 3. 如何确保所提交的数据真实存在，且可以查证？
期刊如何确保数据论文质量？	1. 在对数据论文的审核中，需要考虑哪些因素？例如数据期刊的初审要考虑哪些问题。 2. 数据论文中的内容应该遵循怎样的原则？例如数据期刊的复审要考虑哪些问题。 3. 可以利用哪些第三方过滤机制来协助审核数据论文中所提到的数据，剔除不符合规范的数据？例如怎样的知识库才是适合数据存储的位置。
如何保护利益相关人的各项权益？	1. 读者的权益，例如他们怎样可以找到最新版本的数据？ 2. 作者的权益，例如他人如何承认作者提交数据的贡献？其他人应该如何使用作者提供的数据？ 3. 其他权益，例如就保护试验对象而言，应该和不应该披露哪些信息？

表 3 学术期刊与数据期刊的对比

	学术期刊	数据期刊
对象	学术论文	数据集
受众	科研工作者、教师、学生	科研工作者、教师、学生、公众
特点	文章的专业性、观点新颖性、方法的科学性、结论的可靠性	数据收集过程的技术严谨性、数据完整性以及与现有共同标准的匹配度
实质	思想交流和传播	数据访问和引用

如果从“作者提交数据论文后,期刊要确保数据论文质量,并且在发布时保护利益相关人的各项权益”这个思路来考虑数据期刊,则有研究问题三项:

问题 1:作者如何提交数据论文?

问题 2:期刊如何确保数据论文质量?

问题 3:如何保护利益相关人的各项权益?

4 数据期刊提交规范

4.1 数据期刊要求作者提交数据描述符

数据期刊以文章的形式发布有科学价值的数据描述,也就是数据描述符(Data Descriptor,即“数据论文”)。*SD* 中的数据集主要是指计算或策划数据,以及通过实验或观察产生的数据,包括“技术验证”(Technical Validation)和“用法说明”(Usage Notes)部分,另外在 2014 年 11 月 13 日也开始发布社会科学领域的数据[35]。数据论文作为传统研究出版物的补充,可用于描述已经在其他出版物中分析到的数据集,或用来描述那些独立的数据集,遵循一定的数据标准、计算机可读、可检索[8]。这些描述符不应当包括本质上新的科学方法的描述,也不应当包括新的科学假说的测试。

数据描述符,能够将传统的叙述性内容与结构化描述的研究数据结合在一起,提供一个新框架,用于数据共享和重用,以期促进科学探索。

4.2 数据描述符的内容格式

在 *SD* 发表的数据论文具有六项特点:(1)迅速:快速地审查和出版;(2)严格:具有同行评审;(3)开放:数据描述符免费提供给所有公众,而作者具有署名权;(4)可见:强化浏览和搜索使得数据描述符更容易被发现;(5)相互链接:数据描述符链接到出版集团其他刊物上的相关文章,例如 *SD* 的数据论文链接到 Nature 出版商的其他学术论文;(6)全球化,例如基于 nature. com 网站与世界各地的媒体报道保持联系等。为了能够产生更多的可引用、可发现、可解释、可重用和可再生的科学数据,所提交的数据论文应当对数据描述进行结构化和规范化,其手稿需要包含以下要素并符合对应的格式要求[36],如表 4。

表 4 数据描述符格式要求

元素	要求
标题	包括空格在内最多 110 个字符
摘要	不包括引用最多 170 个字
方法	不限制长度
数据记录	不限制长度
技术验证	不限制长度
用法说明	不限制长度
图	一般每篇手稿不超过 3 个
表	一般每篇手稿不超过 10 个

相比一般学术期刊上对科研论文有研究目的、对象、方法、结果等的描述和记录,在数据期刊上的数据描述符则对于数据本身进行更结构化地描述,有利于信息检索、分析、挖掘以及相关利用。图 2 显示了一篇具体的数据论文《The multilayer temporal network of public transport in Great Britain》[37]。

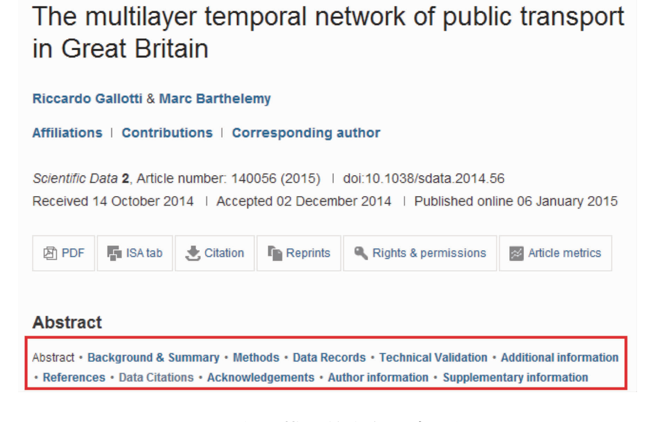


图 2 数据描述符中各要素示例

4.3 作者必须确保所提交的数据真实存在且可以查证

作者将他们的数据集存放在核准的数据知识库中,这是稿件提交过程的一部分,否则稿件不会进入评审过程。如果数据集在提交的时候不对公众开放,作者必须提供安全链接或通行码确保评审

人可以以保密方式访问和评估数据。这是提交数据论文的必要条件，也是发展数据期刊必须保证数据真实存在所提出的必要性政策规范。

5 数据期刊质量规范流程

5.1 对于数据描述符的内容信息是否符合描述规范进行审核

编辑和审稿人会评估作者所选择的用于存储数据的知识库是否合适、数据收集过程的技术严谨性、所得数据集的重用价值、数据描述的完整性以及与现有共同标准的匹配度。在必要时，还会要求作者提供实验质量控制的证据。在需要支持数据的证据时，审稿人可能会要求支持实验的额外证据。

5.2 对于数据描述符所描述的数据及其产生流程的科学性进行专家评审

不同的学科类别会成立专门的编辑委员会来对数据进行辨识判定，大多数数据描述符将由至少一个具有相关实验技术专业知识的科学家和一个数据标准专家进行评审。审核过程中需要考虑到以下几点：(1)实验方法的有效性；(2)第三方使用数据的完整性；(3)数据描述符与数据内容一致连贯；(4)数据能否被开放获取和使用等。

5.3 对于数据描述符所描述的数据具体存储位置进行查验

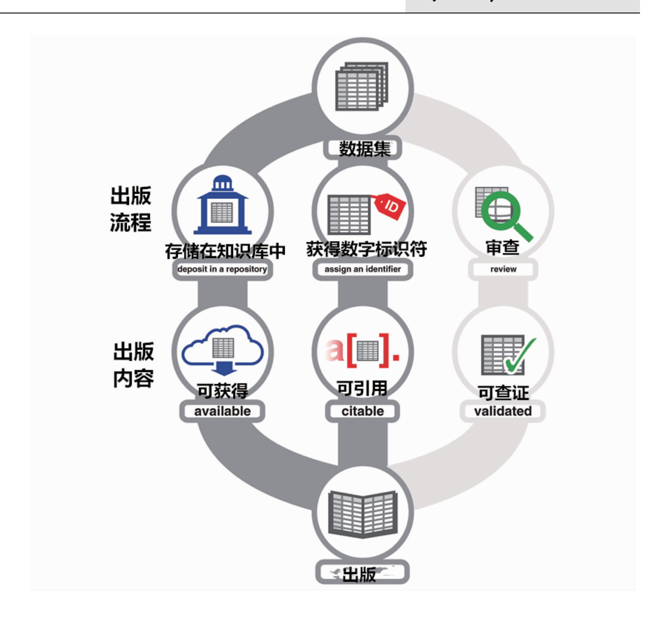
要求将数据描述符存储到规范的数据知识库中，这是对数据质量的一个辅助把关。*SD* 规定所有的数据集应提交给相应的外部知识库，大力支持创建社会各界所公认的数据知识库。如果没有相应的知识库，或者可用的知识库不支持保密的同行评审过程，则要求作者将他们的数据提交到通用的知识库中，如 figshare[38] 或 Dryad[39]。作者还被要求在手稿中提供登录号，便于查找和引用。参见图 3[40]。

良好的数据知识库应当获得科学界的广泛支持和承认，进而能够确保其数据集出版形式的持久性，提供专家管理，落实相关的、社会认可的报告要求，并为提交的数据集提供保密审查和提供稳定的标识符。

6 数据期刊权益规范

6.1 读者的权益

有利于读者利用数据更新找到最新版本的数



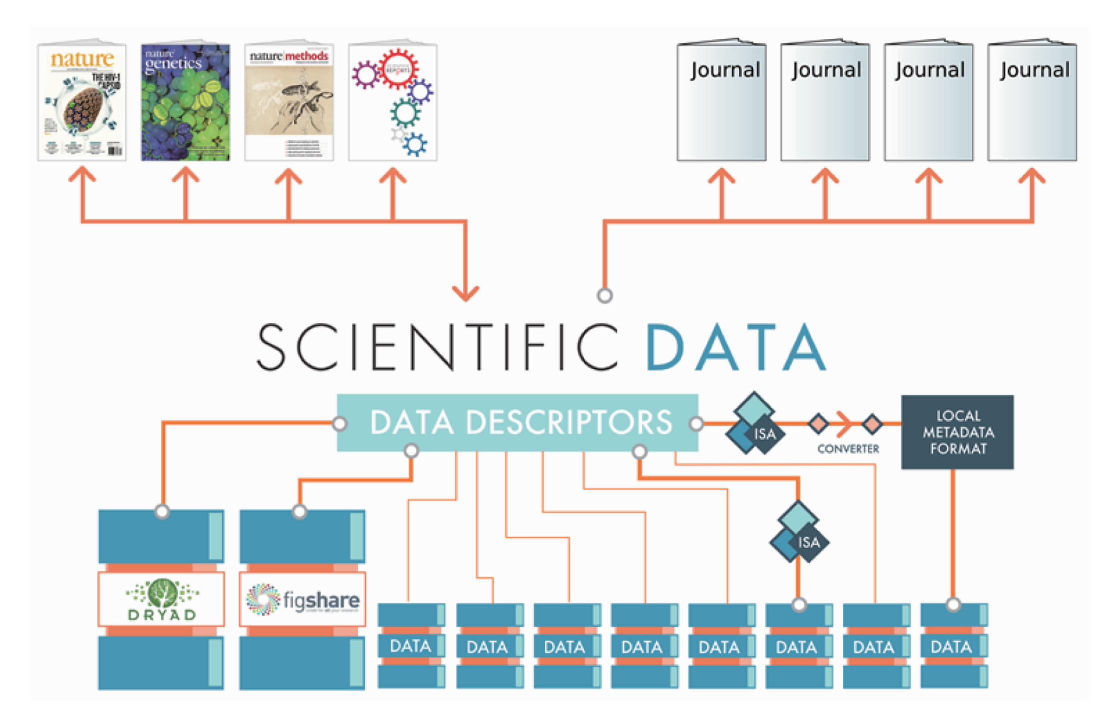
注：如图所示，数据集通常存储在知识库中以便于获取，同时还分配有一个对应标识符，便于引用。一些出版商还将审查数据集以验证[40]。

图 3 数据出版的概要

据。*SD* 要求作者必须承诺在发布数据描述符至少五年内，将他们的数据集保存在其实验室或机构的服务器上，这与英国科研资助机构对数据长期保存的最低要求一致。如果在这段时间里，最初提交数据的知识库出现问题或经验数据丢失，可以要求作者将数据上传到另一个知识库，并刊登更正或更新原有数据描述符。

重要数据集往往会随着时间更新和发展，因此保持发表稳定版本的数据显得尤为重要。区别静态版本的数据与持续演变的数据，并鼓励知识库实行版本控制机制，鼓励作者与这些知识库维护者讨论可用的知识库版本控制机制。*SD* 的做法是将静态版本的数据存入核准的知识库当中（figshare 和 Dryad），同时将持续演变的不同版本存储在具体的项目资源中，这样用户就可以找到最新的数据。这两个版本的数据集都应该在稿件的数据记录部分加以说明。参见图 4[41]。

如果数据期刊能够提供一个可搜索的出版平台（在技术上并不是太难，因为他们已经能够出版“数据期刊”了），科研人员就可以在许多不同的数据存储库中找到高质量的数据集，有利于数据期刊的影响力和公信力。例如 *SD* 就和两大通用知识库 figshare 和 Dryad 合作，进行数据存储和查验，同时还计划使用 ISA（Investigation、Study、Assay）框架与其他知识



注:公开可用的科学数据分布在众多不同的存储库中,因而很难找到相关的数据集(也就是常说的"数据竖井"的问题)。*SD* 推荐存储在通用知识库 figshare 和 Dryad 中[41]。

图 4　*SD* 的出版及其存储数据知识库

库之间开发元数据传输。由此可知,数据期刊既可以依托一个数据存储发布平台,也可以依托已有的数据知识库等,作为检查和存储实际数据的位置。

对数据更新的规范,能够确保读者查找到最新的相关数据,减少重复实验和重复的科学工作,促进科学研究的检验,还有可能与其他数据进行整合,探索新的研究视角。

6.2　作者的权益

作者在提交手稿时需要附带引用权限的声明,规定其他人在使用作者提供的数据时,应该遵循 CC BY 或 CC BY-NC 知识共享许可协议。

例如 *SD* 要求作者在提交手稿时附带有权引用数据的书面保证。在作者引用其他团体或私人来源且未发表的数据或结果时,他们必须保证有适当的权限,并符合创作者的数据许可协议。如果其他人引用数据期刊中的数据论文,则建议用传统的参考文献格式引用数据描述符,如果期刊支持数据引用,则列举出所有引用的数据集。*SD* 认同并遵循数据引文原则的联合声明。作者可以选择两个开放存取创作共用许可[42]:署名 4.0(CC BY4.0)或署名——非商业性使用 4.0(CC BY-NC4.0)。作者需要适当承认由前人出版作品派生的知识或技术贡献,并适当归功于先前公布的数据集的重用,这有助于跟踪数据重用和科学家的数据分享活动。

对引用权限和创作许可的规定不仅能够保证对数据论文的合理使用,更是对作者数据贡献的一种承认,有利于改善当下基于论文的科研评价机制,同时也支持和承认那些提供科研数据的科研人员的贡献。

6.3　其他相关权益

主要涉及到与审稿人和媒体记者的沟通以及有关人体和动物试验者的数据。

数据期刊将所有提交的文稿以及与作者和审稿人的通信视为机密。例如在 *SD* 的规定中,明确要求作者与 *SD* 的通信、评审报告和其他机密材料不得发布于任何网站,或不得未经 *SD* 出版团队事先许可而以其他方式公布。这一做法可以保障独立学术评审不受其他大众媒体舆论压力的影响,对保持学术中立性和独立性具有重要的作用,特别是在提倡开放共享的氛围中,避免少数人员或者利益团体的技巧性控制。

在使用人体和动物为试验对象时,要求作者必

须声明数据集的隐私限制，包括批准该实验的机构及/或许可委员会，涉及到伦理问题、数据和资料的访问以及生物安全的任何细节。由于编辑部本身难以顾及广大科研人员的学术道德问题，因此作为数据期刊应当查证内容以及检验声明，但实际的管理和权责则是在科研人员所属的科研教育机构。

7　数据期刊的政策要素建议

7.1　政策要素内容

综上所述，在考虑我国目前的学术期刊发展实际情况[43]，梳理了数据期刊政策相关重点内容，如表 5 所示。

表 5　数据期刊政策相关重点内容

研究问题	观测问题	政策内容提炼
数据期刊提交规范	提交对象	数据描述符。其中数据集主要是指计算或策划数据，以及通过实验或观察产生的数据
	内容格式要求	标题、摘要、方法、数据记录、技术验证、用法说明
	强制规范	作者必须确保所提交的数据真实可查证，因而作为出版的条件必须将数据集提交到核准的数据知识库中
数据期刊质量规范	数据论文的审核	知识库是否合适、数据收集过程的技术严谨性、所得数据集的重用价值、数据描述的完整性以及与现有共同标准的匹配度
	数据描述符的审核	实验方法的有效性；第三方使用数据的完整性；数据内容与数据描述符一致连贯；数据能否被开放获取和使用
	数据存储	要求作者将他们的数据提交到通用的数据知识库中，如 figshare 或 Dryad，并在手稿中提供登录号，便于检索和引用
数据期刊权益规范	读者的权益	利用数据更新找到最新版本的数据，鼓励作者与数据知识库维护者讨论可用的知识库版本控制机制，鼓励知识库实行版本控制机制
	作者的权益	数据引用及数据许可：作者在提交手稿时需附带引用权限的声明，规定他人在引用数据论文时需遵守 CC BY 或 CC BY-NC
	其他相关权益	涉及人类和动物试验者的数据、伦理问题、数据和资料的访问以及生物安全，作者必须给出授权批准该实验的机构及/或许可委员会，包括方法部分的任何相关细节

需要注意到的是，*SD* 的政策背景不是在中国，在我国实施时，引入这些管理规范能否最大程度地保障我们的相关权益。例如，在表 5 中，我们建议的是 CC BY3.0 而不是 CC BY-NC-ND 4.0 等。目前 CC4.0 是国际协议，虽然开放程度的解释空间较为弹性，但缺乏各国独自进行诠释的途径，我国的 CC 协议是“知识共享 3.0”版本，因此建议在国内环境下使用这个版本中的 CC BY 许可授权；另外，有关数据描述符存储位置，建议存储在我国已经发展较为成熟的机构知识库中，例如中国科学院的机构知识库覆盖全部开放服务的研究所，有着良好的政策支持，能够提供全院科研成果的一站式检索和发现服务，有益于保障国家知识资产的自主保存和永续利用。目前，中国科学院文献情报中心科技信息政策中心也积极支持我国数据知识库的政策调研咨询与研究工作。

7.2　研究局限及未来展望

政策研究需要经过实践历练。数据期刊作为一种新的学术期刊形式，相关的政策内容尚不成熟。本文作者根据实用主义方法论，以文献综述的方式进行分析，以合理竞替方法来提炼政策要素，形成了一个初步框架，如果进一步与实践工作相结合，再考虑到国内科研环境和科研发展的阶段性需求，则需进一步细化和修改政策内容制定更加完善的良好实践指南。

本文通过分析认为数据期刊应该着重数据期刊提交规范、质量管理流程以及保障各方权益的声明。未来研究可以进一步关注数据期刊的质量控制与引用、影响评价以及经营管理等。中国科学院文献情报中心张晓林主任指出：“科技期刊出版者和从业者要关注整个信息环境、学术信息交流体系、科研教育机制的变化。按照有利于科学教育与社会、适应于数字网络时代、充分发挥知识的交流与作用功效去试验、创新和发展，则可能有广阔的生存空间、更好的利润和更有力的影响，变危机为机遇，实现凤凰涅槃”[44]。本文抛砖引玉，进行初步探索。作为促进科技信息开放共享的一种方式，数据期刊逐渐具有促进科研数据传播发展的作用，而其中的政策要素又是实践细节。最重要的是，永不

放弃对读者和作者的关心和服务。

致谢　中国科学院文献情报中心科技信息政策中心组织编译了《数据期刊的编辑出版政策》和《数据期刊的数据政策》两份报告，并在机构知识库(ir. las. ac. cn)上提供开放共享，感谢科技信息政策中心的项目资助。

参考文献

[1] Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities[EB/OL]. [2015-01-21]. http://openaccess. mpg. de/67605/berlin_declaration_engl. pdf.

[2] Open Access policy concerning UNESCO publications [EB/OL]. [2015-01-21]. http://www. unesco. org/new/fileadmin/MULTIMEDIA/HQ/ERI/pdf/oa_policy_rev2. pdf.

[3] Organization for Economic Co-operation and Development (OECD). Principles and Guidelines for Access to Research Data from Public Funding[EB/OL]. [2015-01-21]. http://www. oecd-ilibrary. org/docserver/download/9207043e. pdf? expires = 1421808552&id = id&accname = ocid56017385&checksum = FDEA6A39544B12A9538600E9B9FB0DD3.

[4] The Royal Society. Science as an open enterprise[EB/OL]. [2015-01-21]. https://royalsociety. org/policy/projects/science-public-enterprise/report/.

[5] 顾立平. 科学数据权益分析的基本框架[J]. 图书情报知识，2014(1):34 - 51.

[6] Davidson J, Jones S, Molloy L, et al. Emerging good practice in managing research data and research information within UK Universities[J]. *Procedia Computer Science*,2014, 33:215 - 222.

[7] Hrynaszkiewicz I,新谷洋子. データの再利用を促進するオープンアクセス. オープンデータジャーナル[J]. 情報管理,2014,57(9):629 - 640.

[8] 刘凤红，崔金钟，韩芳桥，等. 数据论文:大数据时代新兴学术论文出版类型探讨[J]. 中国科技期刊研究，2014,25(12):1451 - 1456.

[9] Biodiversity Data Journal[EB/OL]. [2015-01-21]. http://biodiversitydatajournal. com/.

[10] Ecology [EB/OL]. [2015-01-21]. http://www. journalofecology. org/view/0/index. html.

[11] Earth System Science Data[EB/OL]. [2015-01-21]. http://www. earth-system-science-data. net/.

[12] Nature. Availability of Data and Materials[EB/OL]. [2015-01-21]. http://www. nature. com/authors/policies/availability. Html.

[13] Nature Publishing Group. Data publication survey-raw data[EB/OL]. [2015-01-21]. http://figshare. com/articles/Data_publication_survey_raw_data/1234052.

[14] Scientific Data [EB/OL]. [2015-01-21]. http://www. nature. com/sdata/.

[15] 国家农业科学数据共享中心 [EB/OL]. [2015-01-21]. http://www. agridata. cn/.

[16] 国家人口与健康科学数据共享平台 [EB/OL]. [2015-01-21]. http://www. ncmi. cn/1.

[17] 地球系统科学数据共享平台 [EB/OL]. [2015-01-21]. http://www. geodata. cn/Portal/? isCookieChecked = true.

[18] 王卷乐，林海，冉盈盈，等. 面向数据共享的地球系统科学数据分类探讨[J]. 地球科学进展，2014,29(2):265 - 274.

[19] 诸云强，孙九林，冯敏，等. 论地学科研信息化环境[J]. 中国科学院院刊，2013,28(4):501 - 510.

[20] 黎建辉. 中国科学数据建设与共享[EB/OL]. [2015-01-21]. http://ir. las. ac. cn/handle/12502/7444.

[21] Ecological Archives-Data Papers[EB/OL]. [2015-01-21]. http://esapubs. org/archive/.

[22] The Journal of Chemical &Enigineering Data[EB/OL]. [2015-01-21]. http://pubs. acs. org/journal/jceaax.

[23] CODATA's Data Science Journal [EB/OL]. [2015-01-21]. http://www. codata. org/publications/data-science-journal.

[24] Dataset Papers in Science[EB/OL]. [2015-01-21]. http://www. datasets. com.

[25] Earth System Science Data[EB/OL]. [2015-01-21]. http://earth-system-science-data. net/.

[26] Ubiquity Press metajournals [EB/OL]. [2015-01-21]. http://www. metajnl. com.

[27] F1000 Research [EB/OL]. [2015-01-21]. http://f1000research. com.

[28] Giga Science [EB/OL]. [2015-01-21]. http://www. gigasciencejournal. com/.

[29] Geoscience Data Journal [EB/OL]. [2015-01-21]. http://www. geosciencedata. com.

[30] Data in Brief [EB/OL]. [2015-01-21]. http://www. journals. elsevier. com/data-in-brief.

[31] International Journal of Robotics Research DataPapers [EB/OL]. [2015-01-21]. http://ijr. sagepub. com/.

[32] Biodoversity Data Journal[EB/OL]. [2015-01-21]. http://www. pensoft. net/journals/bdj.

[33] 顾立平. 前瞻导论——形塑未来与推动政策[M]. 台北:设计家出版社，2013.

[34] 顾立平. 型人导论——网络用户信息行为与差异化服务策略[M]. 北京:科学技术文献出版社,2013:11 - 13.

[35] Scientific Data now inviting submissions from the social sciences [EB/OL]. [2015-01-21]. http://blogs. nature. com/scientificdata/2014/11/13/scientific-data-now-inviting-submissions-from-the-social-sciences/.

[36] Format of Data Descriptors. [EB/OL]. [2015-01-21]. http://www. nature. com/sdata/for-authors.

[37] Riccardo Gallotti, Marc Barthelemy. The multilayer temporal network of public transport in Great Britain[EB/OL]. [2015-



01-21]. http://www.nature.com/articles/sdata201456.

[38] figshare [EB/OL]. [2015-01-21]. http://figshare.com/.

[39] Dryard [EB/OL]. [2015-01-21]. http://datadryad.org/.

[40] Kratz J, Strasser C. Data publication consensus and controversies[v1; ref status: approved with reservations 1, http://f1000r.es/3ag] *F*1000 *Research*,2014, 3:94 (doi:10.12688/f1000research.3979.1).

[41] Scientific Data to complement and promote public data repositories[EB/OL]. [2015-01-21]. http://blogs.nature.com/scientificdata/2013/07/23/scientific-data-to-complement-and-promote-public-data-repositories.

[42] Open Access[EB/OL]. [2015-01-21]. http://www.nature.com/sdata/about/oa.

[43] 中国科学技术协会. 中国科协科技期刊发展报告(2014)[M]. 北京:中国科学技术出版社, 2014:209 – 225.

[44] 张晓林. 关于进一步推动我国开放获取政策与机制的若干思考[J]. 中国科技期刊研究,2013,24(1):150 – 153.

作者贡献声明:

刘晶晶:论文初稿撰写,英文文献整理、案例分析和总结;

顾立平:研究设计,中文与日文文献整理、政策要素分析,论文终稿修改。

The policy research and analysis of data journals: taking *Scientific Data* as an example

LIU Jingjing[1,2], **KU Liping**[1]

1) National Science Library, Chinese Academy of Science, Beijing 100190, China

2) University of Chinese Academy of Sciences, Beijing 100409, China

Abstract: [**Purpose**] In order to explore the implementation details of good practice, this paper analyzes the policy framework elements of data journals, to benefit the development of Chinese data journals and accelerate the pace of scientific discovery. [**Methodology**] This paper is based on the pragmatic methodology by using literature review with the method of reasonable competing to filter and refine policy elements. [**Findings**] Data journals should focus on their submitted standard (commit object, content's format requirements, ensuring data authentication), and quality control process (contents of the data descriptor, producing process of the data descriptor, audit the data storage repository), and the statement to safeguard the rights and interests of the parties (readers' rights, authors' rights, the rights of scientific research institutions and agencies, and so on). [**Conclusions**] Under the demonstration effect of authoritative journals, policies are formulated to accord with the development of data journals in our country, so as to promote the opening and sharing of research data.

Keywords: Research data; Data sharing; Data journal; *Scientific Data*

(本文责编:梁永霞)